ICASE

A TRUST REGION ALGORITHM FOR UNCONSTRAINED MINIMIZATION:

CONVERGENCE PROPERTIES AND IMPLEMENTATION

Avi Vardi

(NASA-CR-185783) A TRUST REGION ALGORITHM FOR UNCONSTRAINED MINIMIZATION: CONVERGENCE PROPERTIES AND IMPLEMENTATION (ICASE) 33 p

N90-70097

Unclas

00/61 0224388

Report No. 80-35

December 4, 1980

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING

NASA Langley Research Center, Hampton, Virginia

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

A TRUST REGION ALGORITHM FOR UNCONSTRAINED MINIMIZATION: CONVERGENCE PROPERTIES AND IMPLEMENTATION

Avi Vardi
Institute for Computer Applications in Science and Engineering

ABSTRACT

The trust region strategy provides a combination of the steepest descent direction and the quasi-Newton direction which depends on the fit between a quadratic approximation of the function and the function itself.

The BFGS formula is frequently used to update the Hessian of this quadratic approximation; while this update seems the most successful in line search algorithms, it has not been analyzed in conjunction with a trust region strategy. We develop a trust region algorithm using the BFGS update for unconstrained minimization problems; global and superlinear convergence theorems are given.

We demonstrate how to implement this algorithm in a numerically stable way. A computer program based on this algorithm has been written and has performed very satisfactorily on test problems and on a large number of other problems; numerical results are provided.

This research was partially supported under NSF Grants No. ECS-7921279 and No. ENG 76-08749,and under NASA Contract No. NAS1-15810 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

1. Introduction

This paper deals with the problem of minimizing a smooth non-linear function $f: \mathbb{R}^n \to \mathbb{R}$. We derive iterative algorithms for this problem; the algorithm starts at some starting point x^0 and generates a sequence $\{x^k\}$ that generally converges to a local solution of the problem, x^*, such that in a neighborhood of the solution, x^* is a minimizer of the function.

In Section 2 we derive the trust region algorithm; the convergence results of two versions of the algorithm are given in Sections 3 and 4. In Section 5 we discuss the implementation of the algorithm and provide numerical results.

In the development of the algorithm we assume that the first derivative of the function f is available but the algorithm makes no use of second derivatives and instead uses approximations to them.

Notation Convention: Throughout this paper we use superscripts to denote the iteration we are in. Thus s^k is the step at the k^{th} iteration. In order to avoid confusion with powers of matrices we use parentheses in the following way: $(B^k)^j$ denotes the j^{th} power of the matrix B^k where B^k is a matrix associated with the k^{th} iteration.

We are interested in finding a local minimizer for the problem

$$\min_{x \in \mathbb{R}^n} f(x). \qquad (1.1)$$

We will use an iterative method; starting with a point x^0 we will create a finite sequence of points $\{x^i\}$ $i = 1,\dots,\ell$ where x^ℓ satisfies one of the stopping criteria that will be specified. We will assume that f is twice continuously differentiable and that there exists a Lipschitz constant K such that $\|\nabla^2 f(x) - \nabla^2 f(y)\| \le K\|x - y\|$ for all $x,y \in \mathbb{R}^n$ and that f is bounded below.

At a local minimizer x^* the gradient $\nabla f(x^*)$ is zero and the Hessian matrix $\nabla^2 f(x^*)$ is positive semi-definite. We assume that at every local minimizer the Hessian matrix is positive definite. In our algorithms we approximate the Hessian matrix $\nabla^2 f(x)$ by a matrix B. Thus using Taylor's theorem we have for small $\|s\|$

$$f(x+s) \approx q(x+s) \equiv f(x) + \nabla f(x)^T s + \tfrac{1}{2} s^T B s . \qquad (1.2)$$

We will use the BFGS update due to Broyden (1970), Fletcher (1970), Goldfarb (1970), and Shanno (1970) to update B^k at each iteration. If $y^k = \nabla f(x^k + s^k) - \nabla f(x^k)$ and $(y^k)^T s^k > 0$ then

$$B^{k+1} = B^k + \frac{y^k y^{k^T}}{y^{k^T} s^k} - \frac{B^k s^k s^{k^T} B^k}{s^{k^T} B^k s^k} . \qquad (1.3)$$

Otherwise $B^{k+1} = B^k$. If B^k is symmetric and positive definite then so is B^{k+1}; moreover if B^{k+1} is given by (1.3) it satisfies the secant equation $B^{k+1} s^k = y^k$.

A common method of using (1.2) is to compute p^k at each iteration k such that p^k minimizes $q(x^k + s) = f(x^k) + \nabla f(x^k)^T s + \frac{1}{2} s^T B^k s$. The solution $p^k = -(B^k)^{-1} f(x^k)$, the Quasi-Newton (Q.N.) step is a descent direction since $(p^k)^T \nabla f(x^k) < 0$.

Line search techniques are often used to determine μ^k such that $s^k = \mu^k \cdot p^k$ will be an acceptable step. One common strategy imposes the following conditions: (see Wolfe (1969).

At each iteration find μ^k that satisfies

$$\left.\begin{array}{l} f(x^k + \mu^k p^k) \le f(x^k) + \alpha \mu^k \nabla f(x^k)^T p^k \\ \nabla f(x^k + \mu^k p^k)^T p^k \ge \beta \nabla f(x^k)^T p^k \\ \text{where } \alpha, \beta \text{ are constants such that } \alpha < \tfrac{1}{2},\ 0 < \alpha < \beta < 1. \end{array}\right\} \qquad (1.4)$$

Under the assumptions that we made on f, there always exists an interval $[\bar{\mu},\bar{\bar{\mu}}]$ such that if $\mu^k \in [\bar{\mu},\bar{\bar{\mu}}]$, both conditions are satisfied. For a proof see e.g. Dennis and Schnabel (1980).

Powell (1976) showed that if f is strictly convex and when a BFGS update is used to update the Hessian approximations and the two conditions (1.4) are satisfied at each iteration, then $\{x^i\} \to x^*$ and the covergence is superlinear.

2. The Trust Region Model

The quadratic approximation in (1.2) is clearly not valid for all values of s. We will now define a sphere around the current point x^k, $\{x = x^k + s\colon \|s\|_2 \le r^k\}$ and call it the trust region with the interpretation that we trust q only within a radius r^k from x^k. The determination of r^k will be discussed later in this section.

We now want to solve the problem

$$\min_s q(x^k+s) = f(x^k) + \nabla f(x^k)^T s + \tfrac{1}{2} s^T B^k s \qquad \text{s.t. } \|s\|_2 \le r^k. \quad (2.1)$$

In order to obtain a solution to this problem let us rewrite the constraint as $\frac{1}{2} s^T s \le \frac{1}{2}(r^k)^2$. Then we have a convex programming problem for which the Karush-Kuhn-Tucker conditions are sufficient (see, e.g., Avriel). Then if we can find s and λ such that

$$\nabla f(x^k) + B^k s + \lambda s = 0 \tag{2.2}$$

$$\tfrac{1}{2} s^T s - \tfrac{1}{2} r^k \leq 0 \tag{2.3}$$

$$\lambda \cdot (\tfrac{1}{2} s^T s - \tfrac{1}{2} r^k) = 0 \tag{2.4}$$

$$\lambda \geq 0 \tag{2.5}$$

s will be the global solution to (2.1). From (2.2) $s = s(\lambda) = -(B^k + \lambda I)^{-1} \nabla f(x^k)$. If (2.3) is satisfied for $s = s(0)$ then $s(0) = -(B^k)^{-1} \nabla f(x^k)$ is the global solution of (2.1). Otherwise $\tfrac{1}{2}(s(0)^T s(0) - (r^k)^2) > 0$. Clearly for $\lambda >> 0$, $s(\lambda)^T s(\lambda) - (r^k)^2 < 0$ and since $s(\lambda)$ is a continuous function of λ, there exists $\lambda > 0$ with $\tfrac{1}{2}(s(\lambda)^T s(\lambda) - (r^k)^2) = 0$. Such a step once again satisfies (2.2) - (2.5) and therefore is the global solution of (2.1).

The resulting step in the k'th iteration, s^k, associated with λ^k is:

$$\left.\begin{array}{lll} & s^k = s^k(\lambda^k) = -(B^k + \lambda^k I)^{-1} \nabla f(x^k) & \\ \text{where} & \lambda^k = 0 & \text{if } ||s^k(0)|| \leq r^k \\ & \lambda^k > 0 \text{ is s.t. } ||s^k(\lambda^k)|| = r^k & \text{otherwise} \end{array}\right\} \tag{2.6}$$

In the following discussion we omit the superscripts; also ∇f denotes $\nabla f(x^k)$.

<u>Lemma 2.2</u>. The norm of the step $s(\lambda)$ as a function of $\lambda \geq 0$, i.e., $||s(\lambda)||_2 = ||-(B+\lambda I)^{-1} \nabla f||_2$ is a continuous convex function monotonically decreasing to 0 as $\lambda \to \infty$.

<u>Proof</u>. See Vardi (1980).

After a solution to (2.1) is obtained we check whether $f(x+s) < f(x)$. If this is not the case, we will reject the step, and because the quadratic approximation did not prove to be reliable on the sphere with radius r, we will lower the radius and recompute the step.

The following lemma will be needed when analyzing the algorithm.

Lemma 2.3 For any integer $p > 0$, matrix B and $g \in \mathbb{R}^n$, we have that $\lambda^p \cdot (B+\lambda I)^{-p} \cdot g \to g$ as $\lambda \to \infty$

Lemma 2.4. Let

$$s_r \equiv s(\lambda(r)) = -(B+\lambda(r)I)^{-1}\nabla f(x) \tag{2.7}$$

where $\lambda(r)$ is such that $||s(\lambda(r))||_2 = r$ (for $r \leq ||B^{-1}\nabla f(x)||_2$). Then if $\nabla f(x) \neq 0$, there exists $r > 0$ small enough such that $f(x+s_r) < f(x)$.

Proof. Let

$$t(r) = f(x+s_r). \tag{2.8}$$

It is sufficient to show that for all $r > 0$ small enough, $t'(r) < 0$.

$$\frac{\partial t}{\partial r} = \frac{\partial t}{\partial \lambda} \cdot \frac{\partial \lambda}{\partial r} = \nabla f(x-(B+\lambda(r)I)^{-1}\nabla f(x))^T (B+\lambda(r)I)^{-2}\nabla f(x) \cdot \frac{\partial \lambda}{\partial r}.$$

Since $r = \| s_r \|_2$ we find that

$$\frac{\partial r}{\partial \lambda} = -[\nabla f(x)^T (B+\lambda(r)I)^{-2}\nabla f(x)]^{-1/2}[\nabla f(x)^T (B+\lambda(r)I)^{-3}\nabla f(x)]$$

$$= -\frac{\nabla f(x)^T (B+\lambda(r)I)^{-3}\nabla f(x)}{r} < 0. \tag{2.9}$$

To complete the proof we have to show that for $r > 0$ small enough

$$\nabla f\left(x - (B + \lambda(r)I)^{-1}\nabla f(x)\right)^T \left(B + \lambda(r)I\right)^{-2}\nabla f(x) > 0 \ . \qquad (2.10)$$

When $r \to 0$, $\nabla f(x - (B + \lambda(r)I)^{-1}\nabla f(x)) \to \nabla f(x)$ and by using lemma 2.3, $(\lambda(r))^2(B + \lambda(r)I)^{-2}\nabla f(x) \to \nabla f(x)$. Thus (2.10) is satisfied.

The trust region algorithm can be viewed as a combination between the quasi Newton direction and the steepest descent direction. The reason is that when $r >> 0$, reflecting a complete trust in the quadratic approximation, the trust region includes the Q. N. step, $\lambda = 0$ and $s = -B^{-1}\nabla f(x)$. When r is very small, reflecting very little trust in the quadratic approximation, $\lambda >> 0$ and the term λI dominates the term B in the expression $s = -(B + \lambda I)^{-1}\nabla f(x)$. (See Lemma 2.3) Therefore $s \approx -\frac{1}{\lambda}\nabla f(x)$ and this is a short step along the steepest descent direction.

Here is a first sketch of the algorithm.

<u>Algorithm I</u>:

Step 1: **Start** with x^0, r^0, B^0, $k = -1$.

Step 2: $k = k+1$.

Step 3: Find λ^k and s^k such that $||s^k||_2 = ||-(B^k+\lambda^k I)^{-1}\nabla f(x^k)||_2 \leq r^k$ and $\lambda^k(r^k - ||s^k||_2) = 0$.

Step 4: Check if $f(x^k+s^k) < f(x^k)$. If not, reduce r^k and return to Step 3.

Step 5: Check for convergence. If not achieved, continue.

Step 6: Compare $q(x^k+s^k)$ with $f(x^k+s^k)$ and $\nabla f(x^k)+B^k s^k$ with $\nabla f(x^k+s^k)$ and accordingly set r^{k+1}.

Step 7: Compute $y^k = \nabla f(x^k+s^k) - \nabla f(x^k)$ and update $B^{k+1} = \mathrm{BFGS}(B^k, s^k, y^k)$.

Step 8: $x^{k+1} = x^k + s^k$; return to Step 2.

We will analyze two versions of this algorithm. The first one uses similar conditions to (1.4) and is the subject of Section 3. The second version is very close to the way the algorithm is actually implemented and is described in Section 4.

3. Analysis of the First Version of the Algorithm

In order to obtain Algorithm II we first modify the definition of s_r as follows:

$$\left.\begin{array}{ll} s_r = -(B+\lambda I)^{-1}\nabla f(x) \text{ with } \lambda \geq 0 \text{ s.t. } \|s_r\|_2 = r & \\ & r < \|B^{-1}\nabla f(x)\| \\ s_r = -\dfrac{r}{\|B^{-1}\nabla f(x)\|_2} B^{-1}\nabla f(x) \text{ when } r > \|B^{-1}\nabla f(x)\| & \end{array}\right\} \quad (3.1)$$

(The necessity to allow steps that are longer than Q.N. step was first pointed out by Mike Pearlman.)

We also replace the condition $f(x^k + s^k) < f(x^k)$ in Step 4 by two conditions that are similar to the conditions in (1.4).

We accept the step s_r^k if $r = r^k$ satisfies

$$\left.\begin{array}{l} f(x^k + s_r^k) \leq f(x^k) + \alpha \nabla f(x^k)^T s_r^k \\ \nabla f(x^k + s_r^k)^T s_r^k \geq \beta \nabla f(x^k)^T s_r^k \end{array}\right\} \quad (3.2)$$

where s_r^k is as defined in (3.1). The imposition of these conditions will enable us to prove global convergence for this algorithm; unfortunately, it may be impossible to satisfy (3.2) if f is not convex. In order to assure fast local convergence the algorithm will also require taking the Q.N. step $s = -B^{-1}\nabla f(x)$ whenever it satisfies (3.2).

The basic difference between Algorithms I and II is that in II the radius does not play a significant role; the choice of the step at each iteration can be described in terms of λ only. We choose to describe it in terms of the radius to stress the similarities between the two algorithms. In addition, the value of r^k can assist in determining an initial trial value of λ^{k+1} at the (k+1)st iteration.

Algorithm II:

Step 1: Start with x^0, r^0, B^0, $k = -1$; fix α,β such that $\alpha < \frac{1}{2}$, $0 < \alpha < \beta < 1$.

Step 2: $k = k+1$.

Step 3: Find $r = r^k$ such that if s_r^k is computed according to (3.1) both conditions in (3.2) are satisfied. (Try first $r^k = ||(B^k)^{-1}\nabla f(x^k)||$, i.e., check whether the Q. N. step satisfies (3.2).)

Step 4: Check for convergence. If not achieved, continue.

Step 5: Compute $y^k = \nabla f(x^k+s^k) - \nabla f(x^k)$ and update $B^{k+1} = BFGS(B^k,s^k,y^k)$.

Step 6: $x^{k+1} = x^k + s^k$; return to Step 2.

We will show that if Algorithm II is followed then $\liminf_{k\to\infty} \| \nabla f(x^k)\| = 0$. We first show that Algorithm II is well defined.

Lemma 3.1. Assume f is strictly convex. Then there exists an interval $[\bar{r},\bar{\bar{r}}]$ such that for all r in $[\bar{r},\bar{\bar{r}}]$, $f(x+s_r) \leq f(x) + \alpha\nabla f(x)^T s_r$ and $\nabla f(x+s_r)^T s_r > \beta\nabla f(x)^T s_r$. (Again, superscripts are omitted.)

<u>Proof</u>. Let $u(r) = f(x+s_r) - f(x) - \alpha\nabla f(x)^T s_r$. For $r < ||B^{-1}\nabla f(x)||_2$, use the definition of s_r in (3.1) to get

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial\lambda(r)}\,\frac{\partial\lambda(r)}{\partial r} =$$

$$[-\nabla f(x+s_r)^T(B+\lambda(r)I)^{-1}s_r + \alpha\nabla f(x)^T(B+\lambda(r)I)^{-1}s_r]\,\frac{\partial\lambda(r)}{\partial r} \qquad (3.3)$$

By using (2.9) and (3.3)

$$\lim_{r\to 0} u'(r) = \lim_{\substack{r\to 0\\ \lambda(r)\to\infty}} \{-[-\nabla f(x+s_r)^T(B+\lambda(r)I)^{-1}s_r+\alpha\nabla f(x)^T(B+\lambda(r)I)^{-1}s_r]$$

$$\times[\nabla f(x)^T(B+\lambda(r)I)^{-2}\nabla f(x)]^{1/2}/[\nabla f(x)^T(B+\lambda(r)I)^{-3}\nabla f(x)]\}.$$

$$= [\nabla f(x)^T\nabla f(x)]^{1/2}(-1+\alpha)/[\nabla f(x)^T\nabla f(x)] < 0$$

Since $u(0) = 0$, we obtain, for $r > 0$ small enough, $u(r) < 0$. If for all $r > 0$ $u(r) < 0$, using the definition of s_r for $r > ||B^{-1}\nabla f(x)||$ in (3.1), we would get a contradiction to the assumption that f is bounded below.

Let $\bar{\bar{r}}$ be the smallest $r > 0$ such that $u(r) = 0$. Let $v(\theta) = f(x+\theta s_{\bar{\bar{r}}}) - f(x) - \theta\alpha\nabla f(x)^T s_{\bar{\bar{r}}}$. v is convex. Moreover, $v(0) = v(1) = 0$. Hence $v'(1) = \nabla f(x+s_{\bar{\bar{r}}})^T s_{\bar{\bar{r}}} - \alpha\nabla f(x)^T s_{\bar{\bar{r}}} \geq 0$. Thus $\nabla f(x+s_{\bar{\bar{r}}})^T s_{\bar{\bar{r}}} \geq \alpha\nabla f(x)^T s_{\bar{\bar{r}}} > \beta\nabla f(x)^T s_{\bar{\bar{r}}}$. Because $\nabla f(x+s_r)^T s_r - \beta\nabla f(x)^T s_r$ is continuous in r, there exists an $\bar{r} \in (0,\bar{\bar{r}})$ such that for all $r \in [\bar{r},\bar{\bar{r}}]$, $\nabla f(x+s_r)^T s_{\bar{r}} > \beta\nabla f(x)^T s_r$. Because $u(r) < 0$ for all $0 < r < \bar{\bar{r}}$, both conditions are satisfied in $[\bar{r},\bar{\bar{r}}]$.

From now on we assume that at each iteration we can obtain a radius r that satisfies (3.2). The line of proof of global convergence is very close to Powell's (1976, Section 3). Lemmas 3.2 and 3.3 are proved there.

Lemma 3.2. Assume $||\nabla^2 f(x)|| \leq \psi$ for x in the level set $\{x: f(x) \leq f(x^0)\}$. Let $y^k = \nabla f(x^k+s^k) - \nabla f(x^k)$. Then we get the inequality $(\|y^k\|^2)/(s^{k^T} y^k) \leq \psi$.

Lemma 3.3. Assume that $||\nabla^2 f(x)|| \leq \psi$ on $\{x: f(x) \leq f(x^0)\}$. If $B^{k+1} = \mathrm{BFGS}(B^k, s^k, y^k)$ then

$$\prod_{j=1}^{k} \frac{||B^j s^j||^2 \cdot s^{jT} y^j}{(s^{jT} B^j s^j)^2} < (C_4)^k$$

for some constant $C_4 > 0$ and for all $k = 1,2,\ldots$.

Lemma 3.4. For all j

$$\frac{||B^j s^j||^2}{(s^{j^T} B^j s^j)^2} \geq \frac{||\nabla f(x^j)||^2}{[s^{j^T}(-\nabla f(x^j))]^2} . \qquad (3.4)$$

Proof. See Vardi (1980).

Using now Lemma 3.3 and Lemma 3.4, we obtain the inequality

$$\prod_{j=1}^{k} \frac{||\nabla f(x^j)||^2 \cdot s^{j^T} y^j}{[s^{j^T}(-\nabla f(x^j))]^2} < C_4 k \quad \forall\, k=1,2,\ldots . \qquad (3.5)$$

Since at each iteration the β condition in (3.2) is satisified, i. e.,

$$\nabla f(x^j+s^j)^T s^j \geq \beta \nabla f(x^j)^T s^j$$

we have

$$y^{j^T} s^j = [\nabla f(x^j + s_r^j) - \nabla f(x^j)] s^j \geq -(1-\beta)\nabla f(x^j)^T s^j. \tag{3.6}$$

From (3.5) we obtain

$$\prod_{j=1}^{k} \frac{||\nabla f(x^j)||^2}{[s^{j^T}(-\nabla f(x^j))]} < \left(\frac{C_4}{1-\beta}\right)^k = C_5^k \quad \forall\, k=1,2,\ldots\ . \tag{3.7}$$

Theorem 3.5. Assume that for all x in the level set $\{x\colon f(x) \leq f(x^0)\}$ $||\nabla^2 f(x)|| \leq \psi$. Then $\liminf\limits_{k\to\infty} \| \nabla f(x^k)\| = 0$.

Proof. Let θ^j be the angle between s^j and $-\nabla f(x^j)$. From (3.6) we have

$$\prod_{j=1}^{k} \frac{||\nabla f(x^j)||}{||s^j||\cos(\theta^j)} < C_5^k \qquad k=1,2,\ldots\ . \tag{3.8}$$

Recall that the α condition in (3.2) gives at each iteration

$$f(x^j+s_r^j) \leq f(x^j) + \alpha \nabla f(x^j)^T s_r^j .$$

Thus

$$\begin{aligned}\alpha \sum_{j=1}^{k} s_r^{j^T}(-\nabla f(x^j)) &\leq f(x^1) - f(x^2) + f(x^2) - f(x^3) + \ldots \\ &\qquad + f(x^k) - f(x^{k+1}) \\ &\leq f(x^1) - f(x^{k+1}).\end{aligned}$$

f is bounded below and therefore $\sum\limits_{j=1}^{\infty} s_r^{j^T}(-\nabla f(x^j))$ is convergent. Hence $\sum\limits_{j=1}^{\infty}||s_r^j||\cdot||\nabla f(x^j)||\cdot\cos(\theta^j)$ is convergent and $||s_r^j||\cdot||\nabla f(x^j)||\cdot\cos\theta^j$ converges to zero as j tends to infinity.

If we assume now that $||\nabla f(x^j)||$ is bounded away from zero we get a contradiction. We then have $||s_r^j||\cdot\cos(\theta^j) \overset{j\to\infty}{\to} 0$ so that $[||\nabla f(x^j)||]/[||s_r^j||\cdot\cos(\theta^j)] \overset{j\to\infty}{\to} \infty$. This implies that for all sufficiently large j, $[||\nabla f(x^j)||]/[||s_r^j||\cdot\cos(\theta^j)] > 2C_5$. Thus there exists k such that

$$\prod_{j=1}^{k} \frac{||\nabla f(x^j)||}{||s_r^j||\cdot\cos(\theta^j)} > C_5^k$$

which contradicts (3.8). The conclusion is that $\liminf_{j\to\infty} \|\nabla f(x^j)\| = 0$.

<u>Corollary 3.6</u> **When** f is a convex function, if the sequence $\{x^k\}$ has a limit point x^* then x^* is the global minimizer of f. In particular, if f also has a bounded level set, then the hypothesis of the theorem holds and $x^k \to x^*$, the global minimizer.

We now establish the local properties of algorithm II. We will assume, based on our global results, that $\lim x^k = x^*$, a local minimizer, and that $s^k \neq 0$ for all k. The main tools to establish local convergence have been established in a classic paper by Broyden, Dennis and Moré (1974). Lemma 3.7 and Theorem 3.8 are proved there.

Let $M = [\nabla^2 f(x^*)]^{1/2}$. Define the matrix norm $||C||_M = ||MCM||_F$ where $||D|| = ||D||_F = \sqrt{\sum_{ij} d_{ij}^2}$.

We will also use the fact that there exist a constant $\eta > 0$ such that for every n × n matrix A,

$$||A||_F \leq \eta||A||_M. \tag{3.9}$$

Lemma 3.7. For all $u,v \in \mathbb{R}^n$

$$||\nabla f(u) - \nabla f(v) - \nabla^2 f(x^*)(v-u)|| \leq K\sigma(u,v)\cdot||v-u||$$

where $\sigma(u,v) = \max\{||u-x^*||,||v-x^*||\}$. Furthermore, if $\nabla^2 f(x^*)$ is invertible, there is an $\varepsilon > 0$ and $\rho > 0$ such that $\sigma(u,v) \leq \varepsilon$ implies that $\frac{1}{\rho}\,||v-u|| \leq ||\nabla f(v) - \nabla f(u)|| \leq \rho||v-u||$.

Theorem 3.8. There exist $\varepsilon > 0$, $\delta > 0$ such that if $||x-x^*|| < \varepsilon$, $||B^{-1} - \nabla^2 f(x^*)^{-1}||_M < \delta$ and $x^+ = x+s$ for some step s, $B^+ = \mathrm{BFGS}(B,s,y)$ where $y = \nabla f(x+s) - \nabla f(x)$, then B^+ is nonsingular and there exist $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 > 0$ such that

$$||(B^+)^{-1} - \nabla^2 f(x^*)^{-1}||_M \leq [\sqrt{1-\alpha_1\theta^2} + \alpha_2\sigma(x,x^+)]\,||B^{-1}-\nabla^2(x^*)^{-1}||_M + \alpha_3\sigma(x,x^+) \tag{3.10}$$

where $\theta = \dfrac{||[B^{-1}-\nabla^2 f(x^*)^{-1}]y||}{||B^{-1}-\nabla^2 f(x^*)^{-1}||_M\cdot||M^{-1}y||}$.

The following theorem guarantees that in a neighborhood of the solution the Q. N. step $s = -B^{-1}\nabla f(x)$ will satisfy (3.2).

Theorem 3.9. There exist positive ε and δ such that if $||x^0-x^*|| < \varepsilon$ and $||(B^0)^{-1} - \nabla^2 f(x^*)^{-1}||_M < \delta$ then for all k, $r^k = ||(B^k)^{-1}\nabla f(x^k)||$ satisfies both conditions in (3.2). Furthermore, if $x^{k+1} = x^k - (B^k)^{-1}\nabla f(x^k)$ and $B^{k+1} = \mathrm{BFGS}(B^k,s^k,y^k)$, then $||x^{k+1}-x^*|| < \varepsilon$ and $||(B^{k+1})^{-1} - \nabla^2 f(x^*)^{-1}||_M < 2\delta$ for all k. Finally, $||x^{k+1}-x^*|| \leq t||x^k-x^*||$ for $0 < t < 1$ for all k so that $\{x^k\}$ converges to x^* linearly.

Proof. This theorem is similar to theorem 3.4 in Broyden, et al. (1973) and can be found in Vardi (1980).

Lemma 3.10. Let $\{\phi^k\}$, $\{\delta^k\}$ be sequences of nonnegative numbers such that $\phi^{k+1} \leq (1+\delta^k)\phi^k + \delta^k$ for all k and $\sum_{k=1}^{\infty} \delta^k < \infty$. Then $\{\phi^k\}$ converges.

Proof. See Dennis and Moré (1974, lemma 3.3).

Corollary 3.11. Under the conditions of Theorem 3.9, $||(B^k)^{-1} - \nabla^2 f(x^*)^{-1}||_M$ converges.

Proof. From $||x^{k+1} - x^*|| \leq t||x^k - x^*||$ we conclude that

$$\sum_{k=1}^{\infty} ||x^k - x^*|| < \infty. \tag{3.11}$$

Hence from (3.10) and lemma 3.10 the conclusion follows.

Theorem 3.12 will be a step in showing superlinear convergence to x^*. This theorem is proved in Dennis and Moré (1974).

Theorem 3.12. There exist $\varepsilon > 0$ and $\delta > 0$ such that if $\| x^0 - x^* \| < \varepsilon$, $||(B^0)^{-1} - \nabla^2 f(x^*)^{-1}||_M < \delta$ then

$$\lim_{k\to\infty} \frac{||[(B^k)^{-1} - \nabla^2 f(x^*)^{-1}]y^k||}{||y^k||} = 0 \tag{3.12}$$

where $y^k = \nabla f(x^{k+1}) - \nabla f(x^k) \quad \forall\, k = 1,2,\ldots$.

Theorem 3.13. There exist $\varepsilon > 0$, $\delta > 0$ such that if $||x^0-x^*|| < \varepsilon$, $||(B^0)^{-1} - \nabla^2 f(x^*)^{-1}||_M < \delta$ then q-superlinear convergence is achieved, i.e.,

$$\lim_{k\to\infty} \frac{||x^{k+1}-x^*||}{||x^k-x^*||} = 0.$$

Proof. Let ε,δ be as in Theorem 3.9. Observe that

$$\begin{aligned}[B^k-\nabla^2 f(x^*)]s^k = {} & [I-B^k\nabla^2 f(x^*)^{-1}][y^k-\nabla^2 f(x^*)s^k] \\ & -B^k[(B^k)^{-1}-\nabla^2 f(x^*)^{-1}]y^k.\end{aligned}$$

This implies

$$\begin{aligned}\frac{||[B^k-\nabla^2 f(x^*)]s^k||}{||s^k||} \le {} & \frac{||I-B^k\nabla^2 f(x^*)^{-1}||\cdot||y^k-\nabla^2 f(x^*)s^k||}{||s^k||} \\ & + \frac{||B^k||\cdot||[(B^k)^{-1}-\nabla^2 f(x^*)^{-1}]y^k||}{||y^k||}\cdot\frac{||y^k||}{||s^k||} \qquad (3.13)\end{aligned}$$

Lemmas 3.2, 3.7, Theorem 3.12 and the assumptions imply that

$$\lim_{k\to\infty} \frac{||[B^k-\nabla^2 f(x^*)]s^k||}{||s^k||} = 0. \qquad (3.14)$$

ε and δ were chosen in Theorem 3.9 such that for all k, $s^k = -(B^k)^{-1}\nabla f(x^k)$.

Thus

$$[B^k-\nabla^2 f(x^k)]s^k = -\nabla f(x^k)-\nabla^2 f(x^*)s^k = y^k-\nabla^2 f(x^*)s^k - \nabla f(x^{k+1}).$$

Thus lemma 3.7 and the limit in (3.14) imply that

$$\lim_{k\to\infty} \frac{||\nabla f(x^{k+1})||}{||s^k||} = 0. \tag{3.15}$$

Now lemma 3.7 also guarantees that because $\nabla^2 f(x^*)$ is positive definite, if ε is small enough there exist $\rho > 0$ such that

$$||\nabla f(x^{k+1})|| = ||\nabla f(x^{k+1}) - \nabla f(x^*)|| \geq \frac{1}{\rho} ||x^{k+1}-x^*||$$

and therefore

$$\frac{||\nabla f(x^{k+1})||}{||s^k||} = \frac{||\nabla f(x^{k+1})||}{||x^{k+1}-x^k||} \geq \frac{1/\rho||x^{k+1}-x^*||}{||x^{k+1}-x^*||+||x^k-x^*||} = \frac{1}{\rho}\,\frac{\tau^k}{1+\tau^k}$$

where $\tau^k = ||x^{k+1}-x^*||/||x^k-x^*||$. Thus (3.15) implies that $\tau^k/(1+\tau^k)$ converges to zero and hence τ^k converges to zero as desired.

4. Algorithm III and its Local Properties

In order to obtain Algorithm III, we only have to clarify Step 6 in Algorithm I. Step 6 now becomes

$$\left.\begin{aligned}
r^{k+1} &= 2||s^k|| \text{ if } 10\,[\nabla f(x^k)^T s^k + \tfrac{1}{2} s^{k^T} B^k s^k] \geq f(x^k+s^k)-f(x^k) \\
r^{k+1} &= ||s^k|| \text{ if } 0.1[\nabla f(x^k)^T s^k + \tfrac{1}{2} s^{k^T} B^k s^k] > f(x^k+s^k)-f(x^k) \\
&\qquad\qquad > 10[\nabla f(x^k)^T s^k + \tfrac{1}{2} s^{k^T} B^k s^k] \\
r^{k+1} &= \tfrac{1}{2}||s^k|| \text{ if } f(x^k+s^k)-f(x^k) \geq 0.1[\nabla f(x^k)^T s^k + \tfrac{1}{2} s^{k^T} B^k s^k].
\end{aligned}\right\} \tag{4.1}$$

The following lemma guarantees that $r^{k+1} \geq \|s^k\|$ near the solution.

Lemma 4.1. There exist positive ε and δ such that if $||x-x^*|| < \varepsilon$, $||x+s-x^*|| < \varepsilon$ and $||B^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$ then

$$f(x+s)-f(x) < 0.1[\nabla f(x)^T s + \tfrac{1}{2} s^T Bs].$$

Proof.

$$-[\nabla f(x)^T s + \tfrac{1}{2} s^T Bs] = -[\nabla f(x)^T (B+\lambda I)^{-1}(B+\lambda I)s + \tfrac{1}{2} s^T Bs]$$

$$= -[-s^T(B+\lambda I)s + \tfrac{1}{2} s^T Bs] \geq \tfrac{1}{2} s^T Bs \geq \tfrac{1}{2} \frac{||s||^2}{||B^{-1}||_2}.$$

$f(x+s) - f(x) - \nabla f(x)^T s - \frac{1}{2} s^T Bs = \frac{1}{2} s^T(\nabla^2 f(x+\nu s)-B)s$ for some $\nu \in [0,1]$. By using the triangular inequality and the Lipschitz condition we obtain $f(x+s)-f(x)-\nabla f(x)^T s - \frac{1}{2} s^T Bs \leq \frac{1}{2}||s||^2[||B-\nabla^2 f(x^*)||+K(||x-x^*||+||s||)]$. Choose positive ε,δ such that $||x-x^*|| < \varepsilon$, $||x+s-x^*|| < \varepsilon$ and $||B^{-1} - \nabla^2 f(x^*)^{-1}||_M < 2\delta$ imply $0.9/||B^{-1}||_2 \geq ||B-\nabla^2 f(x^*)||+K(||x-x^*||+||s||)$. Now, if $||x-x^*|| < \varepsilon$, $||x+s-x^*|| < \varepsilon$ and $||B^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$,

$$f(x+s)-f(x) = f(x+s)-f(x)-\nabla f(x)^T s - \tfrac{1}{2} s^T Bs + (\nabla f(x)^T s + \tfrac{1}{2} s^T Bs)$$

$$\leq \tfrac{1}{2}||s||^2[||B-\nabla^2 f(x^*)||+K(||x-x^*||+||s||)]+(\nabla f(x)^T s+ \tfrac{1}{2} s^T Bs)$$

$$\leq \tfrac{1}{2}||s||^2 \frac{0.9}{||B^{-1}||_2}+(\nabla f(x)^T s + \tfrac{1}{2} s^T Bs) \leq 0.1(\nabla f(x)^T s + \tfrac{1}{2} s^T Bs).$$

The following theorem guarantees q-superlinear convergence of the algorithm in a neighborhood of the solution.

Theorem 4.2. There exist positive ε and δ such that if $||x^0-x^*|| < \varepsilon$ and $||(B^0)^{-1}-\nabla^2 f(x^*)^{-1}||_M < \delta$ and if Algorithm III is followed, then for all k $||x^k-x^*|| < \varepsilon$ and $||(B^k)^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$. Further, there exist ℓ large enough such that $\lambda^k = 0$ for all $k \geq \ell$. Finally, $||x^{k+1}-x^*|| \leq t||x^k-x^*||$ for $0 < t < 1$.

Proof. By induction.

Let ε_1, δ_1 be the ε and δ from theorem 3.8. Let ε_2, δ_2 be the ε and δ from lemma 4.1. Choose $\gamma > ||\nabla^2 f(x^*)^{-1}||$, $\sigma > ||\nabla^2 f(x^*)||$, and $\xi > 0$. Let ε_3 be such that $||x-x^*|| < \varepsilon_3$ implies that $||\nabla f(x)|| \leq \xi$. Let $\bar{\lambda} = \xi/r^0$, where r^0 is the initial radius. Let $q=\bar{\lambda}/[\bar{\lambda}+1/(2\delta\eta_1+\gamma)]$ where η is defined in (3.9). Let δ_3 be such that $||B^{-1}- \nabla^2 f(x^*)^{-1}||_M < 2\delta_3$ implies $||B^{-1}\nabla f(x^*)|| \leq 1/\sqrt{q}$. Let $t^* \in (\sqrt{q},1)$ and let

$$\delta = \min\{\frac{\delta_1}{2}, \delta_2, \delta_3 \frac{\min\{\frac{1}{3},t^*- \sqrt{q}\}}{3\sigma\eta}\}$$

Let ε_4 be such that

$$\left.\begin{aligned} &\frac{(2\alpha_2\delta + \alpha_3)\varepsilon_4}{1-\max\{\frac{1}{3},t^*\}} \leq \delta \\ &\text{and} \\ &2\delta\eta\sigma + (\gamma+2\eta\delta)K\varepsilon_4 \leq \min\{\tfrac{1}{3},t^*-\sqrt{q}\} \end{aligned}\right\} \qquad (4.2)$$

Let $\varepsilon = \min\{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\}$.

If $\lambda^0 > 0$ then $||s^0|| = ||-(B^0+\lambda^0 I)^{-1}\nabla f(x^0)|| = r^0$. Thus $\lambda^0 \leq ||\nabla f(x^0)||/r^0$ and since $||x^0-x^*|| < \varepsilon_3$, $\lambda^0 \leq \bar{\lambda}$ Note that $||(B^0)^{-1}|| \leq ||[(B^0)^{-1}-\nabla^2 f(x^*)^{-1}]||+||\nabla^2 f(x^*)^{-1}|| \leq 2\eta\delta+\gamma$. Now

$$\begin{aligned}
||x^1-x^*|| &= ||x^0-(B^0+\lambda^0 I)^{-1}\nabla f(x^0)-x^*|| \\
&= ||-(B^0+\lambda^0 I)^{-1}[\nabla f(x^0)-\nabla f(x^*)-\nabla^2 f(x^*)(x^0-x^*)] \\
&\quad + [\nabla^2 f(x^*)^{-1}-(B^0)^{-1}]\nabla^2 f(x^*)(x^0-x^*) \\
&\quad + [(B^0)^{-1}-(B^0+\lambda^0 I)^{-1}]\nabla^2 f(x^*)(x^0-x^*)|| \\
&\le ||(B^0)^{-1}||\ ||\nabla f(x^0)-\nabla f(x^*)-\nabla f(x^*)-\nabla^2 f(x^*)(x^0-x^*)|| \\
&\quad + ||(B^0)^{-1}-\nabla^2 f(x^*)^{-1}||\ ||\nabla^2 f(x^*)||\ ||x^0-x^*|| \\
&\quad + \{\lambda^0\ [\lambda^0+1/||(B^0)^{-1}||]\}||(B^0)^{-1}\nabla^2 f(x^*)||\ ||x^0-x^*||.
\end{aligned}$$

We now use lemma 3.7 and (4.2) to get

$$\begin{aligned}
||x^1-x^*|| &\le [(2\eta\delta+\gamma)\cdot K||x^0-x^*)|| + 2\eta\delta\cdot\sigma + q\frac{1}{\sqrt{q}}]||x^0-x^*|| \\
&\le t^*||x^0-x^*||.
\end{aligned}$$

Thus $||x^1-x^*|| < \varepsilon$.

We next use theorem 3.8 to show (as in theorem 3.4 in Broyden, et al. (1973)) that $||(B^1)^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$. We also notice that lemma 4.1 and Step 6 as set in (4.1) imply that $r^1 \ge ||s^0||$.

Similarly, we can show by induction that if $||x^j-x^*|| < \varepsilon$ and $||(B^j)^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$ for $j=1,2,\ldots,k$ then $||x^{k+1}-x^*|| \le t^*||x^k-x^*|| < \varepsilon$, $||(B^{k+1})^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$ and $r^{k+1} \ge ||s^k||$. Observe that

$$\begin{aligned}
\forall k\ ||s^k|| &= ||x^{k+1}-x^k|| \le ||x^{k+1}-x^*|| + ||x^k-x^*|| \\
&\le 2||x^k-x^*|| \le 2(t^*)^k||x^0-x^*|| \xrightarrow{k\to\infty} 0 \qquad (4.3)
\end{aligned}$$

There must therefore be an ℓ large enough such that $\lambda^{\ell} = 0$. Otherwise, $\lambda^k > 0$ for all k implies that $||s^k|| = r^k \geq r^0$ for all k and this contradicts (4.3).

$$\begin{aligned} ||x^{\ell+1}-x^*|| &= ||x^{\ell}-(B^{\ell})^{-1}\nabla f(x^{\ell})-x^*|| \\ &\leq ||(B^{\ell})^{-1}||\ ||\nabla f(x^{\ell})-\nabla f(x^*)-\nabla^2 f(x^*)(x^{\ell}-x^*)|| \\ &\quad +||(B^{\ell})^{-1}-\nabla^2 f(x^*)^{-1}||\ ||\nabla^2 f(x^*)||\cdot||x^{\ell}-x^*|| \\ &\leq [(2\eta\delta+\gamma)K||x^{\ell}-x^*|| + 2\eta\delta\cdot\sigma]||x^{\ell}-x^*|| \\ &\leq \tfrac{1}{3}||x^{\ell}-x^*|| < \varepsilon. \end{aligned} \qquad (4.4)$$

Also

$$||(B^{\ell+1})^{-1} - \nabla^2 f(x^*)^{-1}||_M < 2\delta.$$

In addition, we have

$$\begin{aligned} r^{\ell+1} \geq ||s^{\ell}|| = ||x^{\ell+1}-x^{\ell}|| &\geq ||x^{\ell}-x^*|| - ||x^{\ell+1}-x^*|| \\ &\geq \tfrac{2}{3}\,\| x^{\ell}-x^*\| . \end{aligned} \qquad (4.5)$$

We will show now that $\lambda^k = 0$ for all $k \geq \ell$. By induction, it is sufficient to prove that $\lambda^{\ell+1} = 0$. Observe first that we can show in a similar way to (4.4) that $||x^{\ell+1}-(B^{\ell+1})^{-1}\nabla f(x^{\ell+1})-x^*|| \leq \frac{1}{3}||x^{\ell+1}-x^*||$. Thus

$$||-(B^{\ell+1})^{-1}\nabla f(x^{\ell+1})|| = ||x^{\ell+1}-(B^{\ell+1})^{-1}\nabla f(x^{\ell+1})-x^{\ell+1}||$$

$$\leq ||x^{\ell+1}-(B^{\ell+1})^{-1}\nabla f(x^{\ell+1})-x^*|| + ||x^{\ell+1}-x^*||$$

$$\leq \tfrac{4}{3}\,\| x^{\ell+1}-x^*\| \leq \tfrac{4}{3}\cdot\tfrac{1}{3}\,\| x^{\ell}-x^*\| \leq \tfrac{2}{3}\,\| x^{\ell}-x^*\| \leq r^{\ell+1}.$$

From the definition of the step (see Algorithm I, step 3), we see that $\lambda^{\ell+1} = 0$.

We complete the proof by induction. If $||x^j-x^*|| < \varepsilon$, $||(B^j)^{-1}-\nabla^2 f(x^*)^{-1}||_M < 2\delta$ for $j=1,2,\dots,k$ and $\lambda^k = 0$, then $||x^{k+1}-x^*|| \le \frac{1}{3}||x^k-x^*|| < \varepsilon$, $||(B^{k+1})^{-1}-\nabla^2 f(x^*)|| < \delta$ and $\lambda^{k+1} = 0$. This completes the proof of the theorem.

Corollary 4.3. The sequence $\{x^k\}$ generated by Algorithm III when $||x^0-x^*|| < \varepsilon$ and $||(B^0)^{-1}-\nabla^2 f(x^*)^{-1}|| < \delta$ converges q-superlinearly.

Proof. Same proof as in the case of Algorithm II. See Theorems 3.12 and 3.13.

5. Implementation

Algorithm I is an iterative one. Steps 3-8 describe what happens in each specific iteration and will be discussed here in detail. Since we confine ourselves in this chapter to one specific iteration, we omit the superscripts.

We want first to comment briefly on scaling. Generally speaking, problem 1.1 is said to be badly scaled if comparable changes in two coordinates of x lead to very different changes in the function f. That can happen when the Hessian matrix of f has a large condition number. When we use the Euclidean norm in (2.1) we have actually assumed that the problem is well scaled. By using a different norm it is possible to have a trust region with an ellipsoid shape rather than a sphere. The solution of (2.1) would then be $s = -(B+\lambda\,\mathrm{diag}(d))^{-1}\nabla f(x)$ where d is some positive vector that reflects the scaling of the problem. The theory of convergence that we presented would still hold for this model. We assume that the variables have been scaled if necessary so that the Euclidean norm is appropriate.

We use the Cholesky decomposition of B, $B = LL^T$. The initial L, L^0 is set to the identity matrix multiplied by the constant $0.1(\|\nabla f(x^0)\|/r^0)^{\frac{1}{2}}$. In each iteration instead of deriving B^{k+1} from B^k, we will derive L^{k+1} from L^k.

Step 3 of the Algorithm:

Assume we are in a new iteration at the point x and have already observed $f(x)$, $\nabla f(x)$. Also available are a lower triangular matrix L and the radius of the trust region. We wish to find a step s as defined in step 3 of the algorithm.

We use the notation

$$s(\lambda) = -(B+\lambda I)^{-1}\nabla f(x) = -(LL^T+\lambda I)^{-1}\nabla f(x).$$

We first compute the Q. N. step $s(0) = -L^{-T}L^{-1}\nabla f(x)$ and hence $\|s(0)\|$ If $||s(0)|| \leq r$ then $\lambda = 0$ and we can go to the next step. If $||s(0)||>r$ we have to find λ such that $||s(\lambda)|| = r$. Define now the function $\phi(\lambda)$ by $\phi(\lambda) = ||s(\lambda)|| - r$; we want to find a zero of this function. From Lemma 2.2 we can conclude that $\phi(\lambda)$ is continuous and convex. $\phi(\lambda)$ decreases montonically to $-r$ and has a unique zero because $\phi(0) > 0$. We know more about the zero, λ^*, of this function; since $\phi(\lambda^*) = ||-(B+\lambda^* I)^{-1}\nabla f(x)||-r = 0$ we have $\lambda^* \leq ||\nabla f(x)||/r$. Also because of the convexity of ϕ, $\lambda^* \geq -\phi(0)/\phi'(0)$. All this information leads to the following iterative process that will converge to a $\bar{\lambda} > 0$ such that $|\phi(\bar{\lambda})| \leq 0.1 \times r$. This iterative process was first suggested by Hebden (1973) and Moré (1978) for Nonlinear Least Squares problems.

Step 1: Let $\ell^0 = -\phi(0)/\phi'(0)$, $u^0 = ||\nabla f(x^0)||/r$.

Let λ^0 be a guess of the solution (which is determined in the previous iteration of the algorithm; see below).

Set $j = 0$.

Step 2: If λ^j does not lie between ℓ^j and u^j,

$\lambda^j = \max\{0.001u^j, (\ell^j \cdot u^j)^{1/2}\}$.

Step 3: Compute $\phi(\lambda^j)$ and if $|\phi(\lambda^j)| < 0.1 \cdot r$, stop.

Step 4: Let

$$\ell^{j+1} = \begin{cases} \ell^i & \text{if } \phi(\lambda^j) < 0 \\ \lambda^j & \text{if } \phi(\lambda^j) > 0 \end{cases}$$

$$u^{j+1} = \begin{cases} \lambda^j & \text{if } \phi(\lambda^j) < 0 \\ u^j & \text{if } \phi(\lambda^j) > 0 \end{cases}$$

Step 5: Compute $\phi'(\lambda^j)$ and set

$$\lambda^{j+1} = \lambda^j - \frac{\phi(\lambda^j)+r}{r} \quad \frac{\phi(\lambda^j)}{\phi'(\lambda^j)} \ ; \ j = j+1;$$

go to step 2.

It may happen rarely that the λ-iterative process does not converge fast enough. We will allow only ten λ-iterations and accept the tenth iteration even if the condition in step 3 is not satisfied.

In our numerical experiments it took 1.6 λ-iterations per x-iteration, on the average, before convergence in step 3 was achieved.

For the different λ's that appear in this λ-iterative process we have to determine $\phi(\lambda)$, $\phi'(\lambda)$. We have

$$\phi(\lambda) = \| s(\lambda) \|_2 - r = [\nabla f(x)^T (B+\lambda I)^{-2} \nabla f(x)]^{\frac{1}{2}} - r \tag{5.1}$$

$$\phi'(\lambda) = - [\nabla f(x)^T (B+\lambda I)^{-3} \nabla f(x)] / [\nabla f(x)^T (B+\lambda I)^{-2} \nabla f(x)]^{1/2} \tag{5.2}$$

$s(\lambda)$ is the solution of the system $(LL^T+\lambda I)s(\lambda) = -\nabla f(x)$, which can also be written as

$$\begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix}^T \begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix} s(\lambda) = \begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix}^T \begin{pmatrix} -L^{-1}\nabla f(x) \\ 0 \end{pmatrix} \tag{5.3}$$

To solve this linear least squares problem we have to find a $(2n \times 2n)$ orthogonal matrix P and an upper triangular R such that

$$P\begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix} = \begin{pmatrix} R \\ 0 \end{pmatrix} \tag{5.4}$$

We note that $L^{-1}\nabla f(x)$ that appears in the right hand side of (5.3) has already been computed. Also notice that because of the special structure of the matrix

$$\begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix}$$

it is possible to apply Givens transformation to this matrix to zero its lower part $(\lambda^{\frac{1}{2}}I)$ without introducing zeroes in the lower triangular part of L^T. We do not have to store P; the computations that have to be done with P will be done while the decomposition is taking place.

Let us partition the matrix P into two $n \times n$ matrices $P = \begin{bmatrix} \bar{P} \\ \bar{\bar{P}} \end{bmatrix}$. Now we have from (5.4)

$$\begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix} = \bar{P}^T R$$

and from (5.3)

$$s(\lambda) = -\left[\begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix}^T \begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix}\right]^{-1} \begin{pmatrix} L^T \\ \lambda^{1/2}I \end{pmatrix}^T \begin{pmatrix} L^{-1}\nabla f(x) \\ 0 \end{pmatrix}$$

$$= -[R^TR]^{-1}R^T\overline{P} \begin{pmatrix} L^{-1}\nabla f(x) \\ 0 \end{pmatrix} = -R^{-1}\overline{P} \begin{pmatrix} L^{-1}\nabla f(x) \\ 0 \end{pmatrix} \tag{5.5}$$

Also from (5.2)

$$\phi'(\lambda) = -\frac{1}{||s(\lambda)||} [s(\lambda)^T(B+\lambda I)^{-1}s(\lambda)]$$

$$= -\frac{1}{||s(\lambda)||} [s(\lambda)^T(R^TR)^{-1}s(\lambda)] = -\|s(\lambda)\| \ \left\| R^{-T} \frac{s(\lambda)}{\| s(\lambda)\|}\right\|. \tag{5.6}$$

Thus in steps 3 and 5 of the λ-iteration process we use (5.5) and (5.6) to compute $\phi(\lambda)$ and $\phi'(\lambda)$.

Steps 4-6 of the Algorithm:

We have three tests that are designed to check three stopping criteria: The f convergence test, ∇f convergence test and x convergence test. In the program the user is asked to specify $\varepsilon_1 \geq 0$, $\varepsilon_2 \geq 0$, $\varepsilon_3 \geq 0$ and funmin - a lower bound on the function.

f convergence test - stop if $f(x)$ - funmin $< \varepsilon_1$

∇f convergence test - stop if $||\nabla f(x)|| < \varepsilon_2$

x convergence test - stop if $||s|| < \varepsilon_3(||x||+1)$.

The user's ε_1, ε_2, ε_3 should depend on the required accuracy in his specific application. For the test problems we use $\varepsilon_1 = 10^{-8}$, $\varepsilon_2 = 10^{-5}$, and $\varepsilon_3 = 10^{-10}$ on our IBM/370.

After s^k is accepted and $x^{k+1} = x^k + s^k$ the program has a series of tests to compare $q(x^{k+1})$ with $f(x^{k+1})$ and $\nabla q(x^{k+1})$ with $\nabla f(x^{k+1})$. The new radius is set to $r^{k+1} = \begin{Bmatrix} \frac{1}{2} \\ 1 \\ 2 \end{Bmatrix} * \| s^k \|$ according to the fit. If $f(x^{k+1}) << q(x^{k+1})$ we recompute the step with $r^k = 2r^k$ before moving to the next iteration. If $f(x_+) > f(x)$ we set $r^k = \frac{1}{2} r^k$ and recompute the step. For complete details for assessing the quadratic model see Vardi (1980).

Step 7 of the Algorithm"

Assume first that s and y have been computed and $s^T y > 0$. We want to obtain $B^+ = \mathrm{BFGS}(B,s,y)$. We have a Cholesky decomposition of $B, B = LL^T$. Let J^+ be such that $J^+ J^{+^T} = B^+$. Dennis and Schabel (1980) derive the BFGS update by

$$J^+ = L + (y - Lv)v^T / v^T v$$

where $v = (y^T s / s^T Bs)^{1/2} L^T s$. Let us take a Q-R decomposition of J^{+^T}, i.e., find Q^+ and L^{+^T} (Q^+ orthogonal, L^{+^T} upper triangular) such that $J^{+^T} = Q^+ L^{+^T}$. Thus $B^+ = J^+ J^{+^T} = L^+ L^{+^T}$ and this gives us the Cholesky decomposition of B^+.

Let $w = (y - Lv)/\| v \|$ and $z = v/\| v \|$. Then $J^{+^T} = L^T + zw^T$. L^T is already an upper triangular matrix and zw^T is a matrix of rank one. Getting a Q-R decomposition of J^{+^T} by using Givens transformations is therefore much cheaper than the $5n^3/2$ operations that a Q-R decompsition require if we made no use of this information. For details see Dennis and Schabel (1980). When $s^T y < 0$ we do not update and take $L^+ = L$.

The program contains a subroutine that obtains a rough estimate of the condition number of B by checking that the ratio between the largest diagonal element and the smallest diagonal element of L is no greater than a prescribed bound. If the ratio is too high the column that contains the smallest diagonal element is changed.

Testing the program:

The program has been tested extensively with test problems that appear in the literature and with problems that were provided by users at Cornell University.

We give the results of four test problems that are often used to test minimization codes. Comparisons of our code, FMIN, with Dennis-Mei's MINOP, Powell's MINFA, Davidon's OCOPTR and Fletcher's VMM01 are made. MINOP and MINFA use, as does our program, an initial radius. Since MINFA's results were reported (by Dennis and Mei (1979)) with one specific r^0 we use this r^0.

Problem 1: Wood's function

$$f(x) = 100(x_2-x_1^2)^2 + (1-x_1)^2 + 90(x_4-x_3^2)^2 + (1-x_3)^2 + 10.1[(x_2-1)^2 + (x_4-1)^2] + 19.8(x_2-1)(x_4-1).$$

Four starting points; $r^0 = 10$.

Problem 2: Rosenbrock's function

$$f(x) = 100\ (x_1^2-x_2)^2 + (1-x_1)^2$$

Five starting points; $r^0 = 3$.

Problem 3: Box's 2-variable function

$$f(x) = \sum_{i=1}^{10} [(e^{-x_1 \frac{i}{10}} - e^{-x_2 \frac{i}{10}}) - (e^{-\frac{i}{10}} - e^{-i})]^2$$

Five starting points; $r^0 = 3$.

Problem 4: Powell's 4-variable function

$$f(x) = (x_1+10x_2)^2 + 5(x_3-x_4)^2 + (x_2-2x_3)^4 + 10(x_1-x_4)^4$$

Three starting points; $r^0 = 3$.

In the next table we report the number of functions and (in parentheses) gradient evaluations for each of the algorithms.

Acknowledgment

My sincere appreciation to M. J. Todd and J. E. Dennis for their guidance and helpful suggestions which have greatly contributed to this work.

Table of Results

Problem Number	Starting Point	FMIN	MINOP	MINFA	OCOPTR	VMM01
1	(-3,-1,-3,-1)	66(44)	71(61)	117(117)	72(50)	129(129)
	(-1.2,1,1.2,1)	81(39)	59(50)	73(73)	83(61)	64(64)
	(-3,1,-3,1)	47(31)	45(35)	94(94)	83(62)	115(115)
	(-1.2,1,-1.2,1)	131(76)	117(103)	119(119)	138(95)	94(94)
2	(-1.2,1)	27(18)	28(25)	37(37)	54(43)	28(28)
	(2,-2)	41(28)	43(42)	39(39)	45(41)	20(20)
	(-3.635,5.621)	38(27)	56(54)	39(39)	56(50)	70(70)
	(6.39,-0.221)	32(19)	61(60)	68(68)	35(32)	74(74)
	(1.489,-2.547)	46(30)	39(36)	43(43)	54(47)	28(28)
3	(5,0)	28(15)	23(20)	20(20)	16(14)	24(24)
	(0,0)	19(14)	18(16)	27(27)	18(17)	21(21)
	(0,20)	29(18)	28(25)	21(21)	19(16)	21(21)
	(2.5,10)	12(4)	8(5)	19(19)	15(14)	16(16)
	(5,20)	26(12)	23(17)	36(36)	26(25)	28(28)
4	(3,-1,0,1)	33(28)	37(35)			
	(-0.1,1.-0.1,1)	33(25)	25(23)	-	-	-
	(-0.6,1,-0.6,1)	35(29)	27(25)		(not available)	

References

Avriel, M. (1976). Nonlinear Programming: Analysis and Methods, Prentice-Hall, Inc.

Broyden, C. G. (1970). The Convergence of Single Rank Quasi-Newton Methods; Math. Comput. 24, pp. 365-382.

Broyden, C. G., Dennis, J. E. and Moré, J. J. (1973). On the Local and Superlinear Convergence of Quasi-Newton Methods; J. Inst. Math. Appl. 12, pp. 223-246.

Cline, A. K., Moler, C. B., Stewart, G. W. and Wilkinson, J. H. (1979). An Estimate for the Condition Number of a Matrix; SIAM J. Numer. Anal. 16, pp. 368-375.

Davidon, W. C. (1959). Variable Metric Method for Minimization; Argonne National Laboratory - 5990.

Dennis, J. E. and Moré, J. J. (1974). A Characterization of Superlinear Convergence and its Application to Quasi-Newton Methods; Math. Comput, pp. 549-560.

Dennis, J. E. and Moré, J. J. (1977). Quasi-Newton Methods, Motivation and Theory; SIAM Review 19, pp. 46-89.

Dennis, J. E. and Mei, H. H. W. (1979). Two New Unconstrained Optimization Algorithms Which Use Function and Gradient Values; J. Optimization Theory Appl. 28, pp. 453-482.

Dennis, J. E. and Schnabel, R. B. (1980). Quasi-Newton Methods for Unconstrained Nonlinear Problems. Manuscript in preparation.

Fletcher, R. and Powell, M. J. D. (1963). A Rapidly Convergent Descent Method for Minimization; Comput. J., pp. 163-168.

Fletcher, R. (1970). A New Approach to Variable Metric Algorithms; Comput. J. 13, pp. 317-322.

Fletcher, R. (1972). An Algorithm for Solving Linearly Constrained Optimization Problems; Math. Programming 2, pp. 133-165.

Gill, P. E., Murray, W. and Pitfield, R. A. (1972). The Implementation of Two Revised Quasi-Newton Algorithms for Unconstrained Optimization; National Physics Laboratory NAC II.

Goldfarb, D. (1970). A Familty of Variable Metric Methods Derived by Variational Means; Math. Comput. 24, pp. 23-26.

Hebden, M. D. (1973). An Algorithm for Minimization Using Exact Second Derivatives; A.E.R.E., Harwell, T.P. 515.

Levenberg, K. (1944). A Method for the Solution of Certain Nonlinear Problems in Least Squares; Quart. Appl. Math. 2, pp. 164-168.

Marquardt, D. W. (1963). An Algorithm for Least Squares Estimation of Nonlinear Parameters; SIAM J. Appl. Math. 11, pp. 431-441.

Moré, J. J. (1978). The Levenberg Marquardt Algorithm: Implementation and Theory; Numerical Analysis, Lecture Notes in Math 630; Editor: Watson; Springer-Verlag.

Ortega, J. M. and Rheinboldt, W. C. (1970). Iterative Solution of Non-Linear Equation in Several Variables; Academic Press.

Pearson, J. D. (1969). Variable Metric Methods of Minimization; Comput. J. 12, pp. 171-178.

Powell, M. J. D. (1970a). A Fortran Subroutine for Unconstrained Minimization Requiring First Derivatives of the Objective Function; A.E.R.E., Harwell, R. 6469.

Powell, M. J. D. (1970b). A Hybrid Method for Nonlinear Equations; Numerical Methods for Nonlinear Algebraic Equations; Editor: Rabinowitz; Gordon and Breach.

Powell, M. J. D. (1971). On the Convergence of the Variable Metric Algorithm; J. Inst. Math. Appl. 7, pp. 21-36.

Powell, M. J. D. (1975). Convergence Properties of a Class of Minimization Algorithms; NLP2; Editors: Meyer and Robinson; Academic Press.

Powell, M. J. D. (1976). Some Global Convergence Properties of a Variable Metric Algorithm for Minimization Without Exact Line Searches. SIAM AMS Proc. 9, pp. 53-72.

Shanno, D. F. (1970). Conditioning of Quasi-Newton Methods for Function Minimization; Math. Comput. 34, pp. 647-656.

Stewart, G. W. (1973). Introduction to Matrix Computation; Academic Press.

Vardi, A. (1980). Trust Region Strategies for Unconstrained and Constrained Minimization. Ph.D. Thesis, School of Operations Research and Industrial Engineering, Cornell University.

Wolfe, P. (1969). Convergence Conditions for Ascent Methods; SIAM Rev. 11, pp. 226-235.